III. *Pedigree Moth-breeding, as a means of verifying certain important Constants in the General Theory of Heredity.* By Francis Galton, F.R.S.

[Read February 2nd, 1887.]

It was suggested by Mr. Merrifield, in answer to my inquiries, that moths, especially those which breed normally twice in the year, would be very suitable subjects for a course of such experimental breedings as I have long desired to establish. My object at the present time is to obtain *data* for the revision and extension of a general theory of simple heredity, on which I have lately published memoirs,* and especially to test that portion of it which relates to Stability of Type. In addition to this, the experiments I propose would elicit incidentally many interesting results, some perhaps quite disconnected with the objects immediately in view.

The merits of moths as subjects of experiment are that the arrangements for breeding them in large numbers occupy comparatively little space, and involve comparatively little cost; their generations succeed each other quickly, and they undergo no change in length or shape of wing, &c., during their brief lives, so that the difficulties elsewhere connected with age and growth disappear; the specimens that are used in the experiments can be afterwards mounted in cases, and be labelled and preserved for future reference.

Mr. Merrifield has very kindly offered to commence the experiments for me, and trusts to be able to continue them for some years. I lay the proposed plan before the Entomological Society in hope of eliciting the suggestions and help of its members before the course of experiments is fairly begun.

---

* "Law of Regression," Journ. Anthrop. Inst., 1885; "Family Likeness in Stature," Proc. Royal Soc., 1886; "Family Likeness in Eye-Colour," Proc. Royal Soc., 1886.

The intention is to start from the brood of a single pair of moths (*Selenia illustraria* is suggested), and to trace the changes of some one characteristic,—say of the wing-length,—during many successive generations, down three parallel but contrasted lines of descent; the broods being reared all along out of doors, and under healthy and perfectly identical conditions. It will be convenient to distinguish these three lines of descent by the letters A, M, and Z. The A line of descent is to be produced by selecting and mating one (or more) pair of the longest-winged males and females that are found in the brood of the original pair. The progeny of this selected pair I will call A i. Out of the males and females in A i. one (or more) pair of the longest-winged will be selected. The progeny of this selected pair will be A ii. Out of the males and females in A ii. one (or more) pair of the longest-winged are again to be selected; their progeny will be A iii. This process will be continued for, it is hoped, at least six generations,—that is, for three years. Every moth in each of the broods is to be preserved, and those which become the parents of the succeeding brood are to be labelled accordingly. It is intended to mount the moths in an orderly series, separating the males from the females, and severally arranging them in the order of their wing-lengths, beginning with the longest and ending with the shortest.

The Z line of descent is to be produced by selecting and mating one (or more) of the shortest-winged males and females out of the original brood. Their progeny will be Z i. Out of these a selection of one (or more) pair of the shortest-winged males and females will be made, whose progeny will be Z. ii., and so on as before. Moths which are small, owing to deformity or obvious unhealthiness, should not be selected for breeding from.

The M line of descent is to be produced by selecting and mating one (or more) pair of medium-winged individuals out of the original brood; their progeny will be M i. Out of these a selection of one (or more) pair of the medium-winged males and females will be made, whose progeny will be M. ii., and so on as before. It will be understood that by medium-winged I mean with reference to the brood in question, and not with reference to the original brood. The use of the M line is to afford a standard whence the divergencies of the A and Z lines in each generation can most suitably be measured. M will be affected, together with A and Z, by all the influences that affect the entire stock, and will therefore exhibit in an unmixed degree the A and Z *peculiarities.* It will easily be understood how important it is to attend to the requirement already laid down that the three lines of descent should be carried on under *identical conditions.*

The broods in the M and Z lines will be mounted in the same way as those in the A line. Therefore for every generation there will be three compartments in one or more trays, each containing (say roughly) 100 moths, and, as there are two generations in a year, the result of each year's breeding will be to fill six compartments.

After the sixth generation or thereabouts has been reached in each of the three lines of descent, it is further desired to proceed conversely, by breeding from medium specimens in each of the three lines, and again from medium specimens of their several broods, and so on until all trace of the A and Z peculiarities shall have disappeared from their respective descendants.

I have spoken of one *or more* pairs, because the moths do not emerge simultaneously, and yet they must be paired soon after they emerge. The ideal pair in the A line would be the very longest-winged male and the very longest-winged female of the entire brood, and in the Z line the very shortest-winged. In practice we must be content with an approximation to this. Two or three separate matings will have to be made between the most suitable of the brood at the time when they require to be paired, but the eggs that are to be preserved and reared will be those of only one or two of the most suitable of the trial pairs. The produce of different pairs ought not to be mixed.

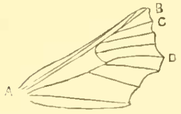
There will be little difficulty about the M pairings, as mediocrities are numerous.

The wing-length is the characteristic with which it is at present proposed to deal, as being more definite and easily measured than others. A similar treatment might be adopted in respect to other characteristics, such as the area of the wing, or the area of the patch of colour in the wings of such moths as may possess it.

A would in this case symbolise large areas, and Z small ones. Or the subject of experiment might be the depth of the general tint. Or, again, it might be the greater or less acuteness of the angle of the wing. Any variable characteristic that exists in both sexes and in all individuals can be treated in this way.

As there is no difficulty in treating the two sexes on equal terms in statistical inquiries by first transmuting all female measurements to their male equivalents [in human stature this is effected by adding one-twelfth or thereabouts; I use the multiplies of 1·08], so probably the different broods of dimorphous moths may be rendered comparable in the same way, and it may be found unnecessary in *Selenia illustraria* to confine the comparisons between spring-brood and spring-broods, and between autumn-brood and autumn-broods respectively. Thus the advantage of double-broodness, in giving two steps a year instead of one, will probably not be lost by experimenting on a dimorphous species.

Measurement.—The wing-length is to be measured from the root of the wing to its tip (that is, from A to B in the fig.). A has to be estimated from the run of the upper and lower margins, and can be determined with fair precision. B is very well-defined in a normal wing. Should, however, the tip be injured, the measurement is to be made from A to C, and stated accordingly. Again, if both B and C are injured, measure from A to D, and state so.



In performing the measurement with a pair of compasses (or other scale) it is necessary to bring the points very close to the moth, else a slight change in the position of the eye will give discordant readings; and if the eye is not moved at all the readings will be sensibly too small. A magnifying-glass is desirable. For measuring living but sluggish moths, I have contrived a pair of compasses that work well and quickly when making the ordinary measurements, and which carry a magnifying-glass. I will not, however, now give a drawing of it, because I want to assure myself of the handiest form of the instrument before doing so. Its principle is this. Imagine a pair of fine-pointed scissors with very long arms,—say five times as long as the blades; the points of the blades represent the compass-points. At the end of one of the arms a scale is pivotted; its free end runs over a pin at the end of the other arm. As the distance between the pivot and the pin is always five times that between the compass-points, the scale is five times the natural size, and can be read off easily to half millimetres, the pin acting as index. In using these compasses the forefinger slightly presses the scale against both the arm on which its free end rests and the pin over which it runs. When the compass is satisfactorily adjusted an increased pressure is sufficient to clamp the scale, and it can then be read off. I have also tried the glasses from one of the tubes of an opera-glass, with a lengthened interval between them, so as to form a microscope of very long focus, say 18 in. This was fixed on a light rod that carried a millimetre scale, set across its free end at a trifle less than 18 in. from the object-glass. On approaching the scale to within half an inch of any small object, that object and the scale are both in fair focus at once, and they are sufficiently far from the eye to render the error of which I spoke of little or no importance.

For the accurate measurement of dead moths I have a much better instrument under construction, in which I use a small microscope with cross wires in the short limb of a pentagraph, and use the long limb both for setting the microscope and for reading off the measurements.

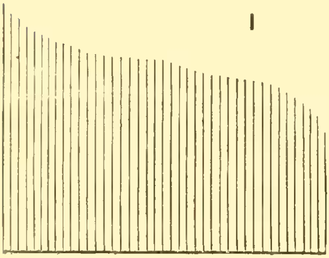
The details of the whole procedure are settled thus far only provisionally, as it is reasonable to hope that much has yet to be gained from the past experience of others, and more by the earlier stages of the experiments themselves, so far as they are new to experience. I have considerable hopes that many persons may feel disposed to work with me, for I am sure it will be accepted as an obvious truth by all, whether they may interest themselves in the technical explanations that follow these remarks or not, that sets of broods of pedigree moths, all of whose direct ancestors, male and female, and all of whose uncles and aunts, great uncles and great aunts, and so on for at least six generations, are preserved in convenient trays for reiterated study, would form a collection of first-class importance for hereditary investigation.

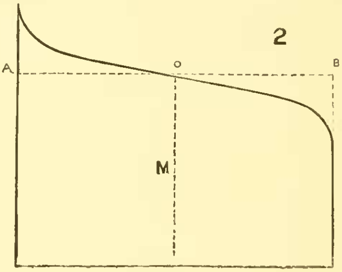
It would be well if in each case of experiment more than one stock could be reared, descended from the same original pair of moths, but in different places, in order that the male and female pupæ may be occasionally interchanged and the moths cross bred, by which the evils of too close interbreeding would be diminished. I sincerely hope that any practical entomologist who may be disposed to further these experiments will communicate with me at 42, Rutland Gate, London, S.W., or with Mr. Merrifield.

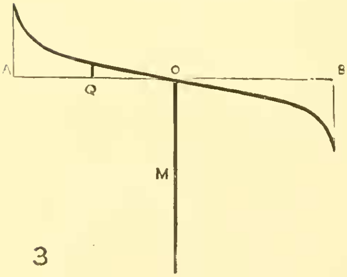
APPENDIX.

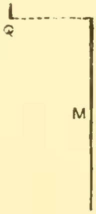
ON THE PROPOSED METHOD OF DEALING WITH THE OBSERVATIONS.

After the results of the experiments have been obtained they will be treated by me in somewhat the same minute and technical way that the data I now possess were treated in the memoirs already referred to. It may perhaps be well to give here a brief and partial account of the guiding principles.


1


2


3


4

When we represent the wing-length of every specimen in a large brood by a corresponding vertical line, and arrange these lines in an orderly manner according to their lengths, and at equal distances apart upon an horizontal base, we shall obtain a figure like that shown in fig. 1. Here, however, only some twenty vertical lines have been drawn, sufficient to indicate what is meant. It is not necessary to make a larger or more minute drawing for the sake of mere explanation. The variability of wing-length, as indicated by the difference between the lengths of the longer and shorter lines, has been purposely exaggerated to make the meaning of the figures more clear.

In fig. 2 these upright lines had been enclosed within a dotted boundary whose vertical sides should lie at an interval of one-half space before the first and beyond the last of the lines respectively, and whose upper portion is a smooth curve drawn with a free hand, touching the tops of the vertical lines; then the vertical lines are supposed to have been rubbed out, and only the contour or "scheme" to remain. This scheme contains, in a most compact form, the measurement of every individual in a brood, or in a population however large. A dotted vertical line, or ordinate, M, is drawn to the curve from a point that bisects the base, and a horizontal line is drawn through the point O in the curve where this ordinate, which is the "median" of the curve, meets it. The median is practically the same as the average, and it is clear from the construction that its value is quite independent both of the width of the scheme and of the number of individuals to which the scheme refers, so long as they are fairly numerous. The horizontal line A O B is the "axis" of the curve; it divides it into symmetrical halves. In fig. 3 the nearer half of the axis A O is itself bisected, and an ordinate Q is drawn to the curve from the point of bisection. Q is what I call the "Quartile" of the curve.

In fig. 4, M and Q are all that remain, and they are all that we are mathematically concerned with. When they and the interval between them are given the whole of the scheme can be calculated; but the interval between them is unimportant for the objects in view. It does not in the least matter on what horizontal scale the scheme is drawn, as the values of Q M and other ordinates at different *fractional* divisions of the axis are independent of the horizontal extension. Q and M are the only values in which we are interested, and it is with these that I work.

It has been abundantly shown by many, from Quetelet onwards, and all my own many statistical inquiries confirm the view, that the curve in the above figure[*] is, or at least tends to be, of as definite a character as an ellipse or any other familiar curve; that it equally admits of mathematical definition, and that it possesses peculiar properties of its own that are of the highest importance to statistical inquirers. Just as an infinite variety of ellipses may be drawn on the axis A B, differing from one another in their extension above and below A B, but otherwise preserving the well-known proportion of an ellipse; so may an infinite variety of *curves of normal variability* be drawn on the axis A B, differing from one another only in the amount of their extension above or below it, as measured conveniently by the length of Q. They will all maintain the proportions shown in the table below, which refers to the middle nine-tenths of the curve. The twentieth part at either end is sure to become irregular. O A is supposed in the Table to be divided into 50 parts, the division at O counting as 0°, and that at A at 50°; then, when Q is taken equal to 100 units, the several ordinates drawn from the principal divisions on the axis are of the lengths shown in the table. For any other value of Q all the tabular values must be changed in uniform proportion to the new value of Q.

	Division.	Length of Ordinate.
	0	0
	10	38
	20	78
Q	25	100
	30	125
	40	190
	45	244
	50	indefinite

As the curve is symmetrical the same measurements apply to either half of it, but in the one half they are made from the axis upwards, in the other half they are made downwards.

If the curve derived from a series of measurements of any variable characteristic is found on trial to conform fairly well with these proportions, it may be assumed that the characteristics in question vary "normally," that is to say, according to the recognised laws of chance, which specify the relative frequency of runs of luck of different lengths. Again, in so far as they vary normally, all the properties of the laws of normal variation may justly be assigned to them. It was by the use of these laws that my deductions were made.

---

[*] The curve actually used by these writers is of another kind and has another signification, but for all that it is the basis of the curve that I employ.

I have exhibited to this Society, as an example of the law of variability, a row of about 100 pods of sweet peas, the produce or brood of a single plant, which I had arranged edgeways, like the vertical lines in fig. 1. Their outline expresses very distinctly the peculiar shape of the curve of variability.

The object of preserving the entire brood of moths is to obtain careful after-measurements from which to deduce the values of M and of Q in each case. When this is done we shall be able to deal with each group in its entirety, and to submit it to mathematical treatment.

The data I have hitherto used in my inquiries were rarely derived from more than three generations, but the condition of statistical constancy in the peculiarities of a population, of which I will again speak, enabled me to extend their scope. They sufficed in this way to lead to many interesting, though perhaps only approximative, results. One is that each parent contributes, on the average, one quarter of the total hereditary peculiarities of the child, each grandparent one-sixteenth, and so on. In other words, that the two parents together contribute one-half, the four grandparents a quarter, the eight great-grandparents one-eighth, and so on, the whole heritage being thus accounted for. But when none of the progenitors besides the two parents are known, their implied peculiarities must be taken into account. They admit of being calculated, and have to be allowed for in the form of an increase to the hereditary contributions of the parents. It is found that each parent should in that case be held to contribute one-third; the difference between one-third and one-quarter (or one-twelfth) being the amount of the implied heritages. It is, however, highly probable from other considerations that, though this simple formula may be closely true for the parents, and nearly true for the grandparents, it may become sensibly and increasingly different for remoter progenitors. It is this that I want to investigate, chiefly through inquiries into Regression. Moreover, all theory concerning the cause and character of Stability of Type, and of much else of high interest in any general view of Heredity, must be based upon the facts of Regression, which such experiments as those proposed can alone, so far as I see, be likely to declare in a trustworthy way.

The laws of simple heredity, as I made them out, involve only five constants. These admit of being separately determined, and they are at the same time connected by an equation that serves to verify their observed values. The equation depends on the fact alluded to, that successive generations of the same population yield identical biological statistics, although each family, or brood, is full of variations, and although the "median" of each characteristic in each brood is on the average *always more mediocre* than the corresponding characteristic in the mean of the two parents. The first of these events, "fraternal variability," increases the variability of the population as a whole, and the latter event, which I call "Regression," decreases it; the two can be shown to counterbalance each other, and give rise to a position of stable equilibrium. The five constants are (1), the Median of the race; (2), the Quartile of the race; (3), the Quartile of the broods of the same parents, *i.e.*, brothers or sisters; (4), the Quartile of broods of a large number of *like* parents, mixed together in a single group; (5), the coefficient of Regression. The laws in which these constants play a part give calculated results that prove to be closely true to observation in the ordinary cases of simple heredity, where there has been no long-continued selection, but it does not at all follow that they will hold true for the descendants of a long succession of widely divergent parents. It is this that I want to test. The point towards which Regression tends cannot, as the history of Evolution shows, be really fixed. Then the vexed question arises whether it varies slowly or by abrupt changes, coincident with changes of organic equilibrium which may be transmitted hereditarily; in other words, with small or large changes of type. Moreover, the values of the Quartile in (3) and (4) cannot be strictly constant, and are probably connected in part with the value of the median, and require a modified treatment by using the geometrical law of error instead of the arithmetical one (Proc. Royal Soc., 1879). Again, the diminution both of fertility and of vitality that accompany wide divergence from racial mediocrity have yet to be measured, by comparing the A and Z broods with the M broods. It was assumed not to vary in the approximative theory of which I spoke.